Prediction of sleep side effects following methylphenidate treatment in ADHD youth

Jae Hyun Yoo\textsuperscript{a}, Vinod Sharma\textsuperscript{b}, Jae-Won Kim\textsuperscript{c,⁎}, Dana L. McMakin\textsuperscript{d}, Soon-Beom Hong\textsuperscript{c}, Andrew Zalesky\textsuperscript{e}, Bung-Nyun Kim\textsuperscript{c}, Neal D. Ryan\textsuperscript{b}

\textsuperscript{a} Department of Psychiatry, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222 Bampo-daro, Seocho-gu, Seoul, 06591, Republic of Korea
\textsuperscript{b} Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, United States
\textsuperscript{c} Division of Child and Adolescent Psychiatry, Department of Neuropsychiatry, Seoul National University Hospital, 101 Daehak-no, Chongno-gu, Seoul, 03080, Republic of Korea
\textsuperscript{d} Department of Psychology, Florida International University, Miami, Florida
\textsuperscript{e} Melbourne Neuropsychiatry Centre, University of Melbourne, Victoria, Australia

ARTICLE INFO

Keywords:
ADHD
Sleep problems
Machine learning
Methylphenidate
Prediction
Side effects

ABSTRACT

Objective: Sleep problems is the most common side effect of methylphenidate (MPH) treatment in ADHD youth and carry potential to negatively impact long-term self-regulatory functioning. This study aimed to examine whether applying machine learning approaches to pre-treatment demographic, clinical questionnaire, environmental, neuropsychological, genetic, and neuroimaging features can predict sleep side effects following MPH administration.

Method: The present study included 83 ADHD subjects as a training dataset. The participants were enrolled in an 8-week, open-label trial of MPH. The Barkley Stimulant Side Effects Rating Scale was used to determine the presence/absence of sleep problems at the 2nd week of treatment. Prediction of sleep side effects were performed with step-wise addition of variables measured at baseline: demographics (age, gender, IQ, height/weight) and clinical variables (ADHD Rating Scale-IV (ADHD-RS) and Disruptive Behavior Disorder rating scale) at stage 1, neuropsychological test (continuous performance test (CPT), Stroop color word test) and genetic/environmental variables (dopamine and norepinephrine receptor gene (DAT1, DRD4, ADRA2A, and SLC6A2) polymorphisms, blood lead, and urine cotinine level) at stage 2, and structural connectivities of frontostriatal circuits at stage 3. Three different machine learning algorithms (Logistic Ridge Regression (LR), support vector machine (SVM), J48) were used for data analysis. Robustness of classifier model was validated in the independent dataset of 36 ADHD subjects.

Results: Classification accuracy of LR was 95.5% (area under receiver operating characteristic curve (AUC) 0.99), followed by SVM (91.0%, AUC 0.85) and J48 (90.0%, AUC 0.87) at stage 3 for predicting sleep problems. The inattention symptoms of ADHD-RS, CPT response time variability, the DAT1, ADRA2A DraI, and SLC6A2 A-3081T polymorphisms, and the structural connectivities between frontal and striatal brain regions were identified as the most differentiating subset of features. Validation analysis achieved accuracy of 86.1% (AUC 0.92) at stage 3 with J48.

Conclusions: Our results provide preliminary support to the combination of multimodal classifier, in particular, neuroimaging features, as an informative method that can assist in predicting MPH side effects in ADHD.

1. Introduction

Attention deficit hyperactivity disorder (ADHD) is a neurodevelopmental disorder characterized by symptoms of inattention, hyperactivity, and impulsivity. Methylphenidate (MPH) is the most frequently prescribed first-line therapeutic agent for ADHD, and it is reportedly effective in approximately 70% of children with the disorder (Santosh and Taylor, 2000). MPH is generally well tolerated by the patients and severe adverse events are rare. A recent meta-analysis also demonstrated that MPH has favorable tolerability and good efficacy in...
child and adolescents and could be recommended as first-choice medica-
tion (Cortese et al., 2018). However, MPH is associated with a dis-
tinctive pattern of relatively common, less severe, but often impairing
side effects that vary from one individual to another as in most psy-
chiatric conditions (Cortese et al., 2013; Vitiello, 2008). From a clinical
perspective, the goal of the initial titration of MPH treatment would be
to establish the best therapeutic response at the lowest dose with the
fewest side effects. Although there has been a considerable degree of
interest in the predictors of therapeutic response to MPH on the
grounds of optimizing individual treatment, limited work has been
done in depth examining the predictors of MPH side effects. Given the
importance of minimizing side effects in establishing the optimal
therapeutic dose of MPH, it could be clinically useful to identify any
biological/cognitive characteristics within the patients that determine
the occurrence of adverse events.

In the current study, we aim to identify predictors of the occurrence of
sleep problems, which is one of the most common side effects of MPH
therapy in ADHD youth (Feldman and Reiff, 2014). Sleep side effects
are important for several reasons that are directly relevant to opti-
mizing clinical care. First, considerable data (both correlational and
causal) indicates that sleep loss negatively impacts daytime attention
and mood in youth (Baum et al., 2013; Beebe, 2011; Beebe et al., 2009,
2010). As such, sleep side effects may be notably counterproductive to
optimal treatment response (Lee et al., 2011), emerge more frequently
in stimulant-naive group (Wigal et al., 2012) and a reason of dis-
continuation (Wigal et al., 2006). Second, several studies indicate that
sleep problems precede the onset of mood disorders (e.g., Gregory et al.,
2009), raising concern that this side effect could increase vulnerability
to comorbid problems with mood in ADHD youth. Finally, increasing
evidence suggests that sleep plays a critical role in aspects of brain
development and learning (Dahl, 2007; Ednick et al., 2009; Tononi and
Cirelli, 2006), with notable interplay between sleep and neural circuitry
involved in attentional and emotional control (e.g. fronto-striatal cir-
cu
try) (Beebe et al., 2009; Yoo et al., 2007). These converging evi-
dences suggest that sleep side effect followed by MPH treatment carry
potential to negatively impact long-term self-regulatory functioning
(Stein, 1999; Van der Heijden et al., 2007).

Shared features of neural systems governing attention and sleep/
wake regulation may partially explain high rates of sleep problems as
side effects of MPH treatment. The presumed therapeutic action of MPH
involves the dopaminergic and noradrenergic neurotransmitter systems
in prefrontal and striatal regions (Wilens, 2008). In particular, MPH
increases dopamine and norepinephrine concentrations in the pre-
frontal cortex. Increases in dopamine and norepinephrine by MPH are
thought to increase wakefulness and lead to enhanced performance on
tasks requiring vigilance and mental awareness, as well as to produce
decreased need for sleep or sleep disturbance (Huang et al., 2011;
Wisler et al., 2001). Although the exact mechanism is unclear, many of
the neurotransmitter systems (e.g., dopamine, norepinephrine, ser-
otonin) and brain regions (prefrontal and striatal [caudate, putamen])
involved in the regulation of sleep/wakefulness (Lazarus et al., 2013;
Vetrivelan et al., 2010) are also related to the regulation of attention/
arousal and pharmacological mechanisms of psychostimulants
(Owens et al., 2013). As such, indices and correlates of functioning in
key neural systems (e.g. fronto-striatal) involved in both sleep/wake
and attention regulation could help to identify those youth at highest
risk for sleep side effects following MPH use.

Dopaminergic and noradrenergic genes have been implicated in
modifying the therapeutic response to MPH (Wilens, 2008). The
dopamine transporter gene (DAT1) and the dopamine D4 receptor gene
(DRD4) are the most extensively studied (Arnst et al., 2011; 
Froehlich et al., 2010). The alpha-2A adrenergic receptor gene
(ADRA2A) and the norepinephrine transporter gene (SLC6A2) have also
been the focus of recent studies (Mick and Faraone, 2008). Candidate
gene predictors of MPH side effects have been reported. For instance,
Stein and colleagues demonstrated that side effects varied as a function
of the DAT1 polymorphism (Stein et al., 2005). McGough and collea-
gues found that an increased risk of irritability was associated with
dopamine gene variants (McGough et al., 2006). However, to date,
there is no study examining objective biological markers that can pre-
dict sleep problems as a side effect of MPH treatment in ADHD.

In addition to genetic factors in ADHD, it is possible that environ-
mental risk factors (e.g., lead, nicotine) and interplay between genetic
predisposition and environmental exposure mediate neurotransmitters
and neurotransmitter systems related to action of MPH (Nigg et al., 2010).
As for neuropsychological endophenotypes of ADHD, the continuous
performance test (CPT) and the Stroop color word test (SCWT) have
been employed to measure neurocognitive functioning and its changes
with MPH treatment in ADHD (Kebrir et al., 2009; Kim et al., 2013).
Thus, these variables may have clinical utility as predictors of treatment
response in ADHD. With regard to neuroimaging measures, diffusion
tensor imaging (DTI) has emerged as a powerful technique for
searching clinically relevant biomarkers or measuring response to
therapy in psychiatric disorders (van Ewijk et al., 2012). DTI research
into ADHD has yielded support to the frontal-striatal model of the dis-
order (Casey et al., 2007; van Ewijk et al., 2012).

Machine learning is an area of artificial intelligence concerned with
the construction and study of systems that can learn from data
(Orru et al., 2012). Recent evidence indicates that the application of
machine learning classification techniques to psychiatric data may
allow prediction of treatment response or side effects at the individual
level (Orru et al., 2012). It is hoped that these methods could inform
and assist clinicians to make more effective clinical decisions prior to
treatment and would lead to fewer unsuccessful trials. To our knowl-
edge, there has been no study that applied machine learning ap-
proaches to predict side effects of MPH treatment in ADHD.

In this study we applied machine learning approaches using pre-
treatment demographic, clinical questionnaire, environmental, neu-
ropsychological, genetic, and neuroimaging information (or features) to
predict the presence or absence of MPH side effects in ADHD youth.
Based on the similarities in neurotransmitter systems and neural cir-
cuity involved in both attention and sleep regulation, we hypothesized
that the biological/cognitive correlates of dopamine/norepinephrine
neurotransmitter systems and prefrontal-striatal neurocircuits in ADHD
would show significant predictive potential for differentiating between
the patients who will develop and not develop sleep problems as a side
effect to MPH administration.

2. Methods

The present study included 83 ADHD subjects (9.5 ± 2.6 years, 65
boys) recruited from the Seoul National University Hospital in Korea.
ADHD was diagnosed according to DSM-IV criteria using the Kiddie-
Schedule for Affective Disorders and Schizophrenia-Present and
Lifetime Version (K-SADS-PL) (Kauffman et al., 1997). ADHD patients
with an intelligence quotient (IQ) below 70, a past or an ongoing his-
tory of either tic disorder, obsessive compulsive disorder, language
disorder, learning disorder, convulsive disorder, pervasive develop-
mental disorder, schizophrenia, bipolar disorder, or brain damage,
a past history of taking stimulants or atomoxetine longer than 6 months,
or a recent history of taking stimulants or atomoxetine over the last 4
weeks were excluded from the study (Hong et al., 2014). The study
protocol was approved by the institutional review board for human
subjects at the Seoul National University Hospital. Detailed information
about the study was given to parents and children, and written in-
formed consents were obtained prior to study entry. At baseline, the
parents completed the ADHD Rating Scale-IV (ADHD-RS) (DuPaul et al.,
1998) and Disruptive Behavior Disorder rating scale (DBD) (Silva et al.,
2005), and the participants undertook the CPT (Greenberg and
Waldman, 1993), SCWT (Golden, 1978), genetic/environmental
testing, and DTI scans.

The participants were enrolled in an 8-week, open-label trial of
MPH. Initial doses of MPH were maintained for 2 weeks, and the doses were adjusted at the 2nd and the 4th week of treatment. The doses were titrated upward until sufficient therapeutic effects were achieved, on the basis of the subjects’ and the parents’ reports of symptom improvement and adverse effects, and then the doses were maintained for the remainder of the 8 weeks. After enrollment at baseline, our study involved four visits at the 2nd, 4th, 6th and 8th weeks. At each visit, we interviewed both the participants and their parents using the Barkley Stimulant Side Effects Rating Scale (Barkley et al., 1990) to determine the presence/absence of side effects. This is a 17-item scale of commonly reported adverse events associated with MPH treatment in ADHD. The “insomnia or trouble sleeping” question was used to determine the presence or absence of sleep problems. If the answers from the participants and their parents were incongruent, we thoroughly examined the reason why one of them reported sleep problems while the other didn’t. After discussion, the interviewer, a well-trained child and adolescent psychiatrist, decided the occurrence of sleep problems if difficulty in falling asleep were at least one day per week after starting MPH treatment. Determination of the sleep side effects was conducted at the 2nd week of the treatment to examine innate susceptibility to sleep side effects at initial dose of methylphenidate whose effects could be similar across participants. In addition, all subjects were retained at the 2nd week without reporting considerable suffering from side effects and deciding to drop-out from the current study.

Genomic DNA was extracted from whole blood lymphocytes using a G-DEX®II Genomic DNA Extraction Kit (Intron, Korea). The DRD4 exon III VNTR polymorphism and the 40-base pair VNTR polymorphism located in the 3’-UTR of DAT1 were genotyped, as previously described (Hong et al., 2012). For the ADR2A and SLC6A2, the detection of a single nucleotide polymorphism was based upon analysis of primer extension products generated from previously amplified genomic DNA, using a chip-based matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry platform (Sequenom, California, USA). The ADR2A (MspI and DraI) and SLC6A2 (G1287A and A-3081T) polymorphisms were genotyped as previously described (Hong et al., 2012). As for environmental factors, we measured blood lead and urine cotinine based on the evidence of our prior ADHD research (Cho et al., 2010). For the lead measurement, a volume of 5 ml of venous blood was collected from each child in metal-free tubes and samples were assayed using previously described methods (Kim et al., 2010). We used urine cotinine as a biomarker for environmental tobacco smoke exposure and it was measured using cotinine direct ELISA kits (BioQuant, San Diego, CA, USA), as previously described (Cho et al., 2013).

The image acquisition and processing implemented herein was based on standard protocols and methods, and is identical to our recent analysis performed in the same cohort (Hong et al., 2014). In brief, for each individual, we seeded streamlines throughout all of white matter and reconstructed the connectome using Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002). The network-based statistic (NBS) (Zalesky et al., 2012, HYPERLINK “http://www.nitrc.org/projects/nbs/”) was used to identify regional brain networks located in the 3

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Demographic, clinical, and neuropsychological characteristics, genotype frequencies, and lead and cotinine levels of the ADHD participants at baseline.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training dataset</td>
<td>Independent dataset</td>
</tr>
<tr>
<td>(n = 83)</td>
<td>(n = 36)</td>
</tr>
<tr>
<td><strong>Age</strong>, mean (SD) years</td>
<td>9.5 (2.6)</td>
</tr>
<tr>
<td><strong>Female</strong>, n (%)</td>
<td>18 (21.7%)</td>
</tr>
<tr>
<td><strong>IQ</strong>, mean (SD)</td>
<td>107 (14)</td>
</tr>
<tr>
<td><strong>Handedness (right)</strong>, n (%)</td>
<td>74 (90.2%)</td>
</tr>
<tr>
<td><strong>CPT</strong>, mean (SD)</td>
<td></td>
</tr>
<tr>
<td><strong>Omission errors</strong></td>
<td>65.7 (20.8)</td>
</tr>
<tr>
<td><strong>Commission errors</strong></td>
<td>64.2 (16.9)</td>
</tr>
<tr>
<td><strong>Response time variability</strong></td>
<td>63.4 (17.6)</td>
</tr>
<tr>
<td><strong>SCWT</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Word test</strong></td>
<td>45.2 (11.0)</td>
</tr>
<tr>
<td><strong>Color test</strong></td>
<td>45.0 (10.5)</td>
</tr>
<tr>
<td><strong>Color-Word test</strong></td>
<td>46.5 (11.7)</td>
</tr>
<tr>
<td><strong>Interference</strong></td>
<td>53.6 (11.3)</td>
</tr>
<tr>
<td><strong>ADHD-RS, mean (SD)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Inattention</strong></td>
<td>15.1 (5.7)</td>
</tr>
<tr>
<td><strong>Hyperactivity-impulsivity</strong></td>
<td>11.0 (5.9)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>26.1 (10.5)</td>
</tr>
<tr>
<td><strong>ADHD subtypes, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>44 (53.0%)</td>
</tr>
<tr>
<td><strong>Inattentive</strong></td>
<td>32 (38.6%)</td>
</tr>
<tr>
<td><strong>Hyperactive-impulsive</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Not otherwise specified</strong></td>
<td>6 (7.2%)</td>
</tr>
<tr>
<td><strong>Comorbid disorders, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Oppositional defiant disorder</strong></td>
<td>16 (19.3%)</td>
</tr>
<tr>
<td><strong>Anxiety disorder</strong></td>
<td>2 (2.4%)</td>
</tr>
<tr>
<td><strong>Genotype</strong></td>
<td></td>
</tr>
<tr>
<td><strong>DAT1</strong>, n (%)</td>
<td></td>
</tr>
<tr>
<td>With 10/10</td>
<td>66 (80.5%)</td>
</tr>
<tr>
<td>Without 10/10</td>
<td>16 (19.5%)</td>
</tr>
<tr>
<td><strong>DRD4</strong>, n (%)</td>
<td></td>
</tr>
<tr>
<td>With 4/4</td>
<td>42 (51.2%)</td>
</tr>
<tr>
<td>Without 4/4</td>
<td>40 (48.8%)</td>
</tr>
<tr>
<td><strong>ADRA2A</strong>, n (%)</td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>38 (46.3%)</td>
</tr>
<tr>
<td>G/C+C/C</td>
<td>44 (53.7%)</td>
</tr>
<tr>
<td><strong>ADRA2A Draz</strong>, n (%)</td>
<td></td>
</tr>
<tr>
<td>C/C</td>
<td>22 (26.8%)</td>
</tr>
<tr>
<td>C/T+T/T</td>
<td>60 (73.2%)</td>
</tr>
<tr>
<td><strong>SLC6A2 G1287A</strong>, n (%)</td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>34 (41.5%)</td>
</tr>
<tr>
<td>G/A+G/A</td>
<td>48 (58.5%)</td>
</tr>
<tr>
<td><strong>SLC6A2 A-3081T</strong>, n (%)</td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>29 (35.4%)</td>
</tr>
<tr>
<td>A/T+T/T</td>
<td>53 (64.6%)</td>
</tr>
<tr>
<td><strong>Environmental measure</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Lead</strong> (µg/dL), mean (SD)</td>
<td>1.5 (0.4)</td>
</tr>
<tr>
<td><strong>Cotinine</strong> (µg/g), mean (SD)</td>
<td>0.7 (1.3)</td>
</tr>
</tbody>
</table>

ADHD, attention deficit hyperactivity disorder; ADHD-RS, ADHD rating scale; ADRA2A, alpha-2A adrenergic receptor gene; CPT, continuous performance test; DAT1, dopamine transporter gene; DRD4, dopamine D4 receptor gene; SCWT, Stroop color word test; SLC6A2, norepinephrine transporter gene. * ADHD (n = 82).
were included for the dataset. We used Wrapper subset evaluation method which evaluates attribute sets by using a learning scheme and identifies features that optimize the prediction performance (Kohavi and John, 1997) Wrapper subset evaluation method was used for feature evaluation and forward greedy hill-climbing augmented with a backtracking algorithm to search and select the subset of feature space. To avoid the overfitting and to increase the generalization of results, Wrapper subset evaluation method was applied with ten-fold cross validation. This gave list of features along with how many times they were chosen (this varies from 0% to 100% time) out of ten folds. Features which were chosen in each fold of ten-fold carried a selection weight of 100% while features which were never selected had a selection weight of 0%. The final feature subset used in this paper was chosen from the feature space generated from each fold of ten-fold cross validation. Generated feature space for this paper mostly included the features which consistently performed well in each fold of ten-fold (30% or more) cross validation stage. Following this method made the results less prone to overfitting and more generalizable to new instances. Features selected have a computational importance with or without clinical relevance or importance; thus, an experienced clinician should decide how to interpret the selected features.

We applied three different machine learning algorithms: support vector machines (SVM), a decision tree algorithm (J48), and Logistic Ridge Regression (a regression-based approach which handles multicollinearity well) with nested ten-fold cross validation to compare the performance of various algorithms. Researchers recommended that it is better to use nested ten-fold cross validation for evaluating the performance of classifiers in case of small sample sizes (e.g., a sample size less than 250) (Hawkins et al., 2003). We chose SVM which is commonly used for modeling complex nonlinear hypothesis spaces when sample sizes are small. We used sequential minimal optimization (SMO) algorithm for training the SVM classifier (Keerthi et al., 2001). We used SVM with second order polynomial kernel and tuned the model for optimum cost parameter from 1 to 100 (details are available upon request). J48 is an implementation of C4.5 decision tree algorithm (Quinlan, 1996), in WEKA. J48 decision tree gives an outcome model which is easy to interpret and applicable in clinical practice. Classification accuracy and area under receiver operating characteristic (ROC) curve (AUC) of these algorithms were compared to find the best classifier for sleep side effects. Based on the cost and burden of obtaining various measures, we set up three stages (see below) and examined whether classification accuracy and AUC could be improved by addition of the measures at each stage. We think that examining tests in tranches based on increasing cost provides data to say whether or not different groups of tests are worth the additional cost.

### Table 2

<table>
<thead>
<tr>
<th>Support Vector Machine</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>J48 Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>Logistic Ridge Regression</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 1</td>
<td>89.7%</td>
<td>40.0%</td>
<td>97.1%</td>
<td>0.83</td>
<td>85.9%</td>
<td>20.0%</td>
<td>95.6%</td>
<td>0.76</td>
<td>85.9%</td>
<td>50.0%</td>
<td>91.2%</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Stage 2</td>
<td>92.3%</td>
<td>40.0%</td>
<td>100%</td>
<td>0.87</td>
<td>83.3%</td>
<td>30.0%</td>
<td>91.2%</td>
<td>0.78</td>
<td>92.3%</td>
<td>70.0%</td>
<td>95.6%</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>Stage 3</td>
<td>91.0%</td>
<td>37.5%</td>
<td>98.3%</td>
<td>0.85</td>
<td>90.0%</td>
<td>25.0%</td>
<td>98.4%</td>
<td>0.87</td>
<td>95.5%</td>
<td>100%</td>
<td>94.9%</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Independent dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 1</td>
<td>58.3%</td>
<td>0.0%</td>
<td>100%</td>
<td>0.50</td>
<td>58.3%</td>
<td>0.0%</td>
<td>100%</td>
<td>0.50</td>
<td>58.3%</td>
<td>0.0%</td>
<td>100%</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Stage 2</td>
<td>66.7%</td>
<td>40.0%</td>
<td>85.7%</td>
<td>0.63</td>
<td>72.2%</td>
<td>40.0%</td>
<td>95.2%</td>
<td>0.71</td>
<td>66.7%</td>
<td>40.0%</td>
<td>85.7%</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Stage 3</td>
<td>66.7%</td>
<td>40.0%</td>
<td>85.7%</td>
<td>0.63</td>
<td>86.1%</td>
<td>86.7%</td>
<td>85.7%</td>
<td>0.92</td>
<td>69.4%</td>
<td>46.7%</td>
<td>85.7%</td>
<td>0.70</td>
<td></td>
</tr>
</tbody>
</table>

Stage 1: demographics and clinical information.
Stage 2: stage 1 + neuropsychological/genetic/environmental measures.
Stage 3: stage 2 + neuroimaging measures.
ADHD, attention deficit hyperactivity disorder; MPH, methylphenidate.

Finally, we tested reproducibility and robustness of our classifier models using an independent dataset. From the study titled “ADHD translational research center” (http://clinicaltrials.gov/show/NCT02623114), an independent group consisted of 36 ADHD subjects (8.5 ± 2.5 years, 32 boys) were recruited from the Seoul National University Hospital. All participants fulfilled the same inclusion and exclusion criteria described above. Demographic, clinical, neuropsychological, genetic/environmental and neuroimaging data were collected in the same manner. For the estimation of frontostriatal connectivity in the independent sample, we applied identical preprocessing methods in the previous research of our group (Hong et al., 2014), except NBS approach. Briefly, deterministic streamlines were generated within 80 pair-wise connections of fronto-striatal circuit. Streamlines selected as a key classifier in the training dataset were registered to FA map of each subject, then a tract-averaged FA value was extracted as features representing structural connectivity. After administration of MPH, sleep side effects at the 2nd week were rated from each subject included in the independent dataset. A set of classifiers in each stage was combined in a step-wise manner, then, applied for the prediction of sleep side effects at the 2nd week of treatment using all 3 machine learning algorithms.

### 3. Results

#### 3.1. Demographic and clinical characteristics

The initial mean dosage of MPH was 16.0 ± 6.2 mg per day (range, 10–36 mg per day) in the training data. The mean overall ADHD-RS score decreased from 26.1 ± 10.5 at baseline to 17.9 ± 9.4 at the 2nd week of treatment. Of the DSM-IV subtypes of ADHD, the combined subtype was the most common (53.0%), followed by the inattentive (38.6%) subtype (Table 1). Among the 83 ADHD subjects, 10 of them developed sleep problems during the first 2-week of MPH trial.

In the independent data, initial mean dosage of MPH was 20.9 ± 9.3 mg (range, 5–45 mg per day), and ADHD-RS score at baseline was 26.0 ± 10.8. The most common subtype of ADHD was combined (47.2%), followed by inattentive subtype (22.2%). Sleep side effects emerged to 15 of 36 subjects (41.7%) in the independent sample.

#### 3.2. Prediction of sleep side effects of MPH treatment

In the training dataset, all three machine learning algorithms well predicted the presence/absence of the sleep side effects after MPH treatment at stage 1 (SVM, accuracy 89.7% (AUC 0.83); J48, accuracy...
85.9% (AUC 0.76); Logistic Ridge Regression, accuracy 85.9% (AUC 0.87), Table 2). Escalated AUC were found when combining neuropsychological and genetic/environmental measures at stage 2.

Logistic Ridge Regression classification accuracy was 95.5% (sensitivity 1.00; specificity 0.95; AUC 0.99) at stage 3 for predicting sleep problems at the 2nd week of treatment (Table 2). Wrapper subset evaluation method demonstrated the inattention symptoms of ADHD-RS, CPT response time variability, the DAT1, ADRA2A DraI, and SLC6A2 A-3081T polymorphisms, and the structural connectivities between the left middle frontal gyrus (orbital part) and left caudate, left inferior frontal gyrus (orbital part) and left putamen, right middle frontal gyrus and right putamen, left superior frontal gyrus (orbital part) and right caudate, and left medial orbitofrontal gyrus and right caudate as the most differentiating subset of features (Fig. 1). SVM and J48 classification accuracies at stage 3 were 69.4% (sensitivity 0.47; specificity 0.86; AUC 0.70) and 66.7% (sensitivity 0.40; specificity 0.86; AUC 0.63) accuracy, respectively.

3.3. Performance validation of key subset of features in the independent dataset

At stage 1, demographic and clinical features performed just above the chance level classification accuracy in all three machine learning algorithm (Table 2). However, step-wise combination of the differentiating features enhanced prediction performance at stage 2 and 3. After combination of structural connectivity features, J48 algorithm predicted 86.1% (sensitivity 0.87; specificity 0.86; AUC 0.92) of the sleep side effect at stage 3. Logistic ridge regression and SVM achieved 69.4% (sensitivity 0.47; specificity 0.86; AUC 0.70) and 66.7% (sensitivity 0.40; specificity 0.86; AUC 0.63) accuracy, respectively at stage 3.

4. Discussion

To our knowledge, the present study is the first to apply machine learning approaches using demographic, clinical neuropsychological, genetic, environmental, and neuroimaging data together to predict MPH side effects in ADHD. As hypothesized, the indices and correlates of functioning in key neurotransmitter and neural systems involved in both sleep/wake and attention regulation showed significant predictive potential for sleep side effects following MPH treatment.

By using three different types of machine learning algorithm (SVM, Logistic Ridge Regression, J48), the best performing algorithm, classification accuracy and AUC continued to improve between all stages and was best at stage 3 (Table 2).
Logistic Ridge Regression and J48 decision tree), the presence/absence of sleep side effects after MPH administration was successfully predicted in the present study. In the training dataset, all three algorithms showed superior prediction performance. A step-wise combination of neuropsychological and genetic/environmental measures (stage 2) as well as neuroimaging features (stage 3) further increased the classification accuracy and AUC. Current findings suggest that individuals with sleep side effects following MPH treatment might have differential profile across gene-brain-behavior. Differential treatment responses to MPH by DAT1, DRD4 and ADRA2A Dra1 genetic polymorphisms (Froehlich et al., 2011; Kim et al., 2010; Winsberg and Comings, 1999), CPT performances (Rapport et al., 1986), or frontostriatal disconnec-tivity (Hong et al., 2015) have been found in ADHD subjects. Together with those evidences, current findings may account for why sleep side effects is present only in a subpopulation of ADHD patients.

Key features from training dataset could well predict presence/absence of sleep side effect after MPH treatment in the independent dataset. The best accuracy was achieved by the J48 algorithm, followed by logistic regression and SVM. Decision trees have a strength in incorporating the interaction effects between variables (Zhao and Zhang, 2008), while linear regression model and SVM do not take these into account. Given the evidence that clinical phenotype is inevitably inter-correlated with genetic or neurobiological changes (Prathikanti and Weinberger, 2005), superior performance of J48 may attribute to consideration of interactions among classifiers in the prediction of new dataset.

Sonuga-Barke and colleagues found that side effects related to sleep were not predictable from patients’ demographic and clinical characteristics, such as age, gender, height, weight, and psychiatric comorbidity (Sonuga-Barke et al., 2009). They suggested that these side effects may be predictable from correlates related to underlying mechanisms of action of MPH. Their finding and suggestion are in line with our results, which showed that the classification accuracy and AUC by Logistic Ridge Regression improved with the addition of neuropsychological/genetic/environmental measures to demographics and clinical information (85.9% to 92.3%, 0.87 to 0.92, respectively), and further improved with the addition of neuroimaging measures (92.3% to 95.5%, 0.92 to 0.99, respectively). Overall, our ability to predict sleep side effects using all variables was successful. However, given the current cost of the imaging and genetic studies and the relatively low risk of treatment with MPH, we do not think that this is of substantial immediate clinical utility.

It is important to note that the DAT1, ADRA2A Dra1, and SLC6A2 A-3081T polymorphisms were included in the most differentiating subset of features. These polymorphisms were reported to be associated with treatment response to MPH (Kim et al., 2010; Hong et al., 2012; Park et al., 2013). Our prior research also suggested ADRA2A Dra1 as one of the key features in predicting MPH treatment responders (Kim et al., 2015). Considering prior evidence and pharmacogenetic action of ADRA2A, norepinephrine pathway could play an important role in both eliciting treatment response and adverse effects with regard to MPH administration in ADHD. It has been well recognized that the mode of action of MPH in ADHD treatment is in its blockade of not only the dopamine transporter (DAT) but also the norepinephrine transporter (NET) (Solanto, 1998; Spencer et al., 1996). In the frontal cortex, where DAT density is low and NET density is higher, it is unlikely that blockade of the DAT accounts for the elevation of dopamine levels in this region. Thus, it has been postulated that dopamine is transported by the NET in the frontal cortex, since dopamine has a higher affinity for the NET as compared with its affinity for the DAT (Madras et al., 2005). ADHD medications have one pharmacological effect in common, which is to elevate extracellular levels of dopamine and norepinephrine in the frontal cortex.

In the current study, serum lead and cotinine level were not selected as key predictors of MPH-induced sleep side effects. These environmental toxins have been associated with symptoms of ADHD in a number of literatures (Cho et al., 2013; Froehlich et al., 2009) as well as increased risk for sleep problems (Liu et al., 2015). However, evidence that environmental toxins may increase sensitivity to sleep side effects following MPH treatment has not yet been identified. Although we could not find meaningful relationship between exposure to environmental toxin and sleep side effects in the present study, potential mechanism of neurotoxicity induced by lead and cotinine, and their impact on pharmacologic side effects may need further investigation.

Collectively, we could successfully predict sleep side effects after MPH administration using several biological markers related with ADHD. However, we did not assess co-existent sleep problems among ADHD subjects at baseline. It has been suggested that subjective reports of sleep disturbances also could be present in medication-free children with ADHD (Konofal et al., 2010). Sleep problems in ADHD might be multifactorial (Owens, 2005), however, underlying pathophysiology is still less well understood. Further studies are needed to understand co-occurrence of sleep problems in medication-free ADHD subjects. Meanwhile, pharmacological actions of MPH on dopaminergic and noradrenergic neurotransmitters (Wilens, 2008), and associated brain circuits have been figured out clearly (Lazarus et al., 2013, Vetivelan et al., 2010). Here, we aimed to examine the occurrence of sleep side effects particularly after MPH treatment based on biological underpinning.

The application of our findings to real-world clinical care will require continued research examining the interplay of MPH, sleep side effects, and treatment outcomes in ADHD. However, there are at least two viable clinical approaches that could be used to manage patients who are identified as vulnerable to sleep side effects, supporting the clinical utility of this investigation. Namely, clinicians could 1) track sleep and titrate medications to maximize therapeutic response while minimizing sleep side effects, and/or 2) provide behavioral sleep intervention or prevention strategies that have proven efficacy in psychiatric populations (Haynes et al., 2006; Troxel et al., 2011).

There are several limitations to this study that deserve comment. First, the sample size was small, and occurrence of sleep side effects were only found in 10 among 83 participants in the training data. Feature extraction can be affected by imbalance of classes, which might be resulted in the low sensitivity of prediction models. Our findings should be replicated in a larger sample. Second, this study was conducted at one university center in Korea and, thus, we are unable to make inference with regard to the generalizability across different research centers or ethnic groups for any of the successful predictors found in our results. Third, we were not able to collect objective data of sleep problems such as polysomnography findings of the patients before the start of MPH treatment. As noted above, we only examined the occurrence of sleep side effects after MPH use; pre-existent sleep problems at baseline were not assessed. Fourth, genetic polymorphisms included in our study were selected from several candidate gene association studies which had limited power and sample size. Further studies with genome-wide association design are needed to confirm whether SNPs associated with DAT or NET are involved in sleep side effects particularly after MPH treatment based on biological underpinning.

In summary, the results of this study demonstrate that sleep side effects following MPH treatment in ADHD can be identified at the individual level using a range of biological and cognitive measures including genetic, neuroimaging, and neuropsychological data. From a clinical perspective, our results provide preliminary support to the combination of multimodal classifier, in particular, neuroimaging features, as an informative method that can assist in predicting response and adverse events to pharmacological treatments in ADHD. Prediction of sleep side effects associated with MPH treatment would not only help decide alternative treatment options initially such as using non-
stimulants (e.g., atomoxetine, clonidine), but also reduce unsuccessful trials and make preventive interventions including alarming the risk of sleep side effects, providing behavioral intervention for sleep improvement or prescribing sleep medications (e.g., melatonin). Still, cost of neuroimaging and genetic polymorphism tests are quite high, which limits immediate utilization in the clinic. Further studies that examine more extensive biological/cognitive correlates that may determine the occurrence of sleep problems in medicated ADHD patients are needed to provide a better understanding of the underlying pharmacological mechanisms of MPH and its role on sleep-wake regulation.

Declaring of Competition Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the Basic Science Program through the National Research Foundation of Korea (2010-0002283 to Dr. J.-W. Kim).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.nicl.2019.102030.

References


Cho, S.C., Hong, Y.C., Kim, J.W., et al., 2013. Association between urine cotinine levels, to provide a better understanding of the underlying pharmacological improvement or prescribing sleep medications (e.g., melatonin). Still, cost


Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Yoo, JH; Sharma, V; Kim, J-W; McMakin, DL; Hong, S-B; Zalesky, A; Kim, B-N; Ryan, ND

Title:
Prediction of sleep side effects following methylphenidate treatment in ADHD youth

Date:
2020-01-01

Citation:

Persistent Link:
http://hdl.handle.net/11343/247003

File Description:
published version

License:
CC BY-NC-ND